

National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California



Semantic Web for Earth and Environmental Terminology (SWEET) 2018: Status, Future Development and Community Building

Lewis J. McGibbney Ph.D., B.Sc.
Computer Science for Data Intensive Applications Group (398M)
Instrument Software and Science Data Systems Section (398)



Agenda and Purpose

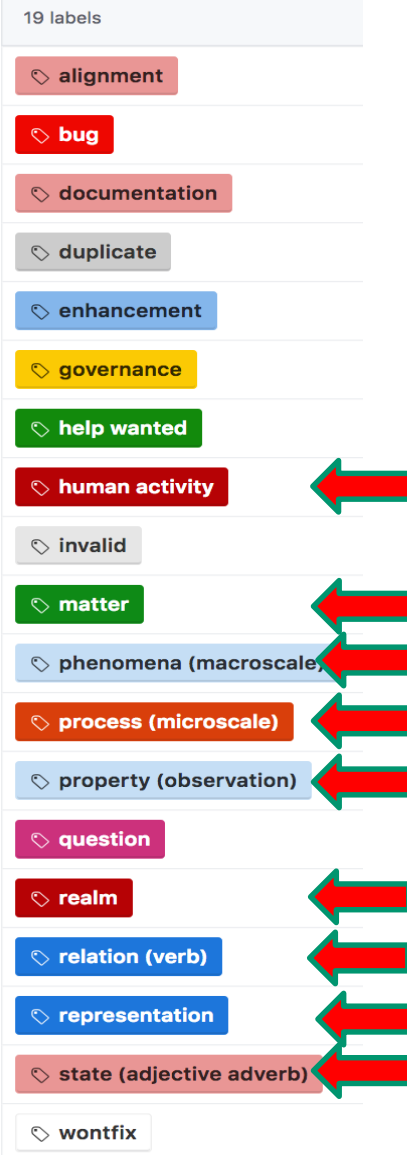
- **Current status of SWEET e.g. What has the community been doing?**
- **Explore the abstract structure of SWEET [1] e.g. What are the primary information topics/categories/clusters and how are they currently extended?**
- **Highlight the development priorities moving forward.**
- **Explore ongoing alignment [2] activities between SWEET and other semantic or linked data resources.**

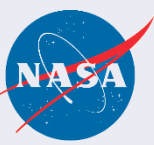


What is SWEET?

- SWEET is a highly modular ontology suite with **6924 concepts** (Classes, Object Property, Data Property and Individuals) in **225 separate ontologies*** covering Earth system science. SWEET is a mid-level ontology and consists of **nine top-level concepts** that can be used as a foundation for domain-specific ontologies that extend these top-level SWEET components.
- SWEET's own domain-specific ontologies, which extend the mid-level ontologies, can provide users interested in further developing a particular domain with a solid set of concepts to get started.
- SWEET ontologies are written in W3C Turtle; the Terse RDF Triple Language and are publicly available under the Apache License v2.0.

* Numbers are accurate as of January 2018







Current status of SWEET

- As of mid 2017 version 3.x of SWEET was recently released under new community development, management and governance by ESIP's Semantic Technologies Committee (STC) [3].
- Development follows a review-then-commit (RTC) policy for source code contributions. The development guidelines can be seen at [4]. These are automatically shown to any contributor upon either creation of a new Github issue or pull request.
- Advances in SWEET since its transition from NASA JPL over to the open source, community-driven management and governance structure now overseen by the ESIP STC include (i) URI transition and governance, transition from OWL to Turtle serialization, linked data dispatch via the ESIP Community Ontology Repository [5], and ongoing alignment activities with existing semantic technology resources such as the OBO Foundry collection, W3C SOSA/SSN, W3C PROV-O, etc.,



URI transition and governance

- The canonical details for this topic can be found at [6]
- Essentially, it boils down to the URI transition from <http://sweet.jpl.nasa.gov/...> to <http://sweetontology.net/...>
- The '#' character has also been substituted for '/' e.g. <http://sweetontology.net/reprDataProduct#Dataset> to <http://sweetontology.net/reprDataProduct/Dataset>
- URI's now no longer contain the file suffix e.g. *.owl. Content negotiation is managed by the hosting platform (COR) depending on the client request. Available serializations include OWL, Turtle, RDF/XML, N-Triples, etc. An example would be <http://sweetontology.net/stateEnergyFlux?format=rdf>



Transition from OWL to Turtle serialization

- SWEET ontologies have recently been transformed to W3C Turtle [7]; the Terse RDF Triple Language.
- We have tools available to ensure that the content of each file is ‘prettified’ e.g. alphabetically ordered, blank nodes use [...] syntax for clearer syntax as supported by the OWL-API.
- This week (ESIP Winter Meeting) Adam Shephard discovered an issue concerning malformed subject URIs for owl:NamedIndividuals that begin with a numeric.
- More information on transition to Turtle can be found at [8]

what is wrong

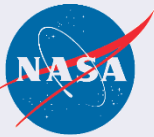
The owl:NamedIndividual URIs for terms that start with a numeric value (seen in the comment preceding the defining triples) is missing the numeric value in the defining triples

Example

```
### http://sweetontology.net/stateTime/1day
stime:day rdf:type owl:NamedIndividual ,
           time:Duration ;
screla:hasUnit units:day ;
screla:hasNumericValue "1.0"^^xsd:double .
```

Should this be:

```
### http://sweetontology.net/stateTime/1day
stime:1day rdf:type owl:NamedIndividual ,
            time:Duration ;
screla:hasUnit units:day ;
screla:hasNumericValue "1.0"^^xsd:double .
```



Linked data dispatch via the ESIP Community Ontology Repository (COR)

- Canonical documentation can be located at [9]
- Full service GUI is available at <http://cor.esipfed.org>
- Core REST API documentation can be located at <http://cor.esipfed.org/ontapi>
- The COR software is being used on several other projects e.g. Marine Metadata Interoperability (MMI) project <https://mmisw.org/>.
- The COR software is maintained at the MMI Github <https://github.com/mmisw> and has an active team of developers.
- Both follow-on sessions below will cover MUCH more information on this
 - GeoSemantics Symposium – 1:30-2:15pm & 3:30-4:15pm (Carlos Rueda, Felimon Gayanilo)
 - COR Developer Workshop - Thursday, January 11 • 3:30pm - 5:00pm <http://sched.co/D6oC> (Carlos Rueda, Beth Huffer, Lewis McGibbney)



Development Priorities (General)

- **ISSUE-20 Annotation of SWEET terms [10]**
 - Very few SWEET terms have human-readable annotations.
 - The definitions of SWEET terms are implied by the axiomatization, which is often incomplete.
 - Every SWEET term should have at least
 - `rdfs:label` (and/or `skos:prefLabel` or `dct:title`) for the name
 - `rdfs:comment` (and/or `skos:definition` or `dct:description`) for a text definition
 - We hope the YAMZ effort will help to move this issue forward. The ESIP-backed source code for YAMZ resides at [11]
 - Also see ISSUE-33 [12] for more information.



Updating OGC Documentation

SWEET is mentioned in a number of places across several OGC Documents, namely

- **Observations and Measurements -**
<http://www.opengeospatial.org/standards/om>
- **Sensor Observation Service -**
<http://www.opengeospatial.org/standards/sos>
- **Sensor Model Language -**
<http://www.opengeospatial.org/standards/sensorml>
- **Sensor Things API -**
<http://www.opengeospatial.org/standards/sensorthings>
- **SWE Common Data Model -**
<http://www.opengeospatial.org/standards/swecommon>

I've recently been granted the ability to continue to OGC through ESIP membership, so I will be actively updating these resources to where SWEET is currently.



Development Priorities – Ontology Alignments

- Ontology alignment, or ontology matching, is the process of determining correspondences between concepts in ontologies. A set of correspondences is also called an alignment.
- Two sub research fields have emerged in recent years in ontology mapping, namely monolingual ontology mapping and cross-lingual ontology mapping. The former refers to the mapping of ontologies in the same natural language, whereas the latter refers to "the process of establishing relationships among ontological resources from two or more independent ontologies where each ontology is labelled in a different natural language".
- Primary resources of information and actual work going on in this field include
 - Ontologymatching.org community [13], and
 - Various output (papers and software) from Ontologymatching.org conferences and journals [14] such as International Workshop on Ontology Matching [15]
- *Ontology alignment is important for furthering the usefulness of SWEET as so much knowledge exists, has been, and is being created by outside experts working in different communities. Our community can only benefit by tapping into and leveraging that expertize and knowledge.*



Development Priorities (Alignments)

- Coordination with **Open Bio Ontologies (OBO)** [16]; preliminary results can be found at [17]. This has resulted in alignments across 15 OBO resources... we need to solicit input from domain experts now to verify alignments before they are formalized.
- Formal alignment between SOSA, SSN and SWEET [18]. Formalizes 7 concepts including **Observation**, **Result**, **Sample**, **Variable**, **Property**, **Sampling** and **Statistical Sample**.
- SWEET dev's have decided that alignment mapping graphs should be limited to one file and are maintained separately from the source and target(s) being mapped.
- Depending on mutability of both SWEET and potential alignment targets, this does increase maintenance... discussion on managing this is required.



Development Priorities – The SWEET Alignment Mapper (SAM)

- **SAM [19] is an attempt to create a generic alignment manager e.g. which can be used for the task of generic ontology alignments, providing a pipeline and results for aligning SWEET with various ontologies.**
- **This is semi-automated but the goal is to learn from a seed of initial curated equivalence axioms using the Agreement Maker Light (AML).**
- **AML [20] is a lightweight ontology matching system specialized on the biomedical domain but applicable to any ontologies. It can be used to generate alignments automatically, as a platform for reviewing alignments, or as an alignment repair system (both automatically and interactively).**
- **Let's look at some alignment outputs...**
- **Preliminary SAM outputs for the Dataset Ontology, PROV-O, SOSA and SSN can also be seen at [21]**



References

1. <https://github.com/esipfed/sweet>
2. <https://github.com/ESIPFed/sweet/tree/master/alignments>
3. http://wiki.esipfed.org/index.php/Semantic_Technologies
4. <https://github.com/ESIPFed/sweet/blob/master/CONTRIBUTING.md>
5. <https://cor.esipfed.org>
6. <https://github.com/ESIPFed/sweet/wiki/SWEET-IRI-Patterns-for-Ontologies-and-Their-Terms>
7. <http://www.w3.org/TR/turtle/>
8. <https://github.com/ESIPFed/sweet-tools>
9. <https://github.com/ESIPFed/sweet/wiki/sweetontology.net>
10. <https://github.com/ESIPFed/sweet/issues/20>
11. <https://github.com/DillonBArevalo/yamz>
12. <https://github.com/ESIPFed/sweet/issues/33>
13. <http://ontologymatching.org/>
14. <http://ontologymatching.org/events.html>
15. <http://om2017.ontologymatching.org/>
16. <http://www.obofoundry.org/>
17. <https://github.com/cmungall/sweet-obo-alignment>
18. <https://github.com/ESIPFed/sweet/blob/master/alignments/sweet-ssn-mapping.ttl>
19. <https://github.com/ESIPFed/sweet-tools/tree/master/sweet-alignment-manager>
20. <https://github.com/AgreementMakerLight/AML-Jar>
21. <https://github.com/ESIPFed/sweet-tools/tree/master/sweet-alignment-manager/alignments>



Acknowledgements

Sponsorship to present this work at the 2018 ESIP Winter Conference was provided by the Jet Propulsion Laboratory, California Institute of Technology, under contract to NASA, the National Aeronautics and Space Administration. This support is highly appreciated.

© 2018 California Institute of Technology. Government sponsorship acknowledged.

Thank you to all previous and current SWEET developers.



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

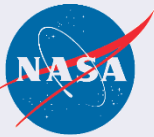
Business Contact

Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, California 91109-8099
Mail Stop : 158-256C
Tel: (+1) (818)-393-7402
Cell: (+1) (626)-487-3476
Fax: (+1) (818)-393-1190
Email: lewis.j.mcgibbney@jpl.nasa.gov
ORCID: orcid.org/0000-0003-2185-928X



Questions and Discussion

- **My intention is utilize any remaining time to actually discuss, or code up some solutions to existing open issues, future development initiatives, alignments, etc.**
- **We have organized several alignment activities for attenees to involve themselves with... the next few slides contain the collaboration locations where we can work on alignments.**



SWEET Github Issue

<https://github.com/ESIPFed/sweet/issues/73>

Existing Preliminary Alignment Output

<https://github.com/ESIPFed/sweet-tools/blob/master/sweet-alignment-manager/alignments/align-sweet-datasetontology.rdf>

Manual Alignment Collaboration Area

<http://bit.ly/2qT7YSc>

(Relevant tab)



SWEET + PROV-O

SWEET Github Issue

<https://github.com/ESIPFed/sweet/issues/28>

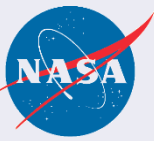
Existing Preliminary Alignment Output

<https://github.com/ESIPFed/sweet-tools/blob/master/sweet-alignment-manager/alignments/align-sweet-prov-o.rdf>

Manual Alignment Collaboration Area

<http://bit.ly/2qT7YSc>

(Relevant tab)



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

SWEET + SSN

SWEET Github Issue

<https://github.com/ESIPFed/sweet/issues/27>

Existing Preliminary Alignment Output

<https://github.com/ESIPFed/sweet-tools/blob/master/sweet-alignment-manager/alignments/align-sweet-ssn.rdf>

Manual Alignment Collaboration Area

<http://bit.ly/2qT7YSc>

(Relevant tab)



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

SWEET + SOSA

SWEET Github Issue

<https://github.com/ESIPFed/sweet/issues/27>

Existing Preliminary Alignment Output

<https://github.com/ESIPFed/sweet-tools/blob/master/sweet-alignment-manager/alignments/align-sweet-sosa.rdf>

Manual Alignment Collaboration Area

<http://bit.ly/2qT7YSc>

(Relevant tab)



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

SWEET + ENVO

SWEET Github Issue

<https://github.com/ESIPFed/sweet/issues/7>

Existing Preliminary Alignment Output

<https://github.com/cmungall/sweet-obo-alignment/blob/master/align-sweet-obo-envo.tsv>

Manual Alignment Collaboration Area

<http://bit.ly/2qT7YSc>

(Relevant tab)



SWEET + PROV-O Alignment

- **My intention is utilize any remaining time to actually discuss, or code up some solutions to existing open issues, future development initiatives, alignments, etc.**
- **We have organized several alignment activities for attendees to involve themselves with... the next few slides contain the collaboration locations where we can work on alignments.**